2020 年全球联网数据库风险分析报告

2020 年 10 月

2020 Global Networked Database Risk Analysis Report

360-CERT COMPUTER EMERGENCY READINESS TEAM

360天枢智库 360 DIPPER RESEARCH

QUAKE

团队介绍

360CERT

360CERT 是政企安全创新中心的尖兵团队，团队致力于维护计算机网络空间安全，是 360 基于 " 协同联动，主动发现，快速响应 " 的指导原则，对全球重要网络安全事件进行快速预警、应急响应的安全协调中心。针对全球重大安全漏洞第一时间启动安全响应流程，发布权威报告，帮助用户进行预防处理，保护用户和互联网安全。

360 天枢智库

360 天枢智库，是中国首家专注于“大安全”研究领域的新型企业智库。天枢智库依托于 360 深耕多年的网络安全实战经验、海量安全大数据、安全创新实践以及完善的专家体系，聚焦网络安全行业发展和大安全问题，看现在，观未来，开展政策、行业、技术等方向的研究。通过与国内外智库平台、高校及科研机构的合作与交流，博采众长，助力网络安全在全行业的积极发展，为助推数字化安全发展建言献策。360 天枢智库秉持科学的、前瞻的、独立的、建设的态度，致力于养成安全领域的思想库和风向标。

360 网络空间测绘系统（Quake）

360 网络空间测绘系统（QUAKE）是 360 网络安全响应中心（360-CERT）自主设计研发的全球网络空间测绘系统，能够对全球 IPv4、IPv6 地址进行持续性探测，实时感知全球网络空间中各类资产并发现其安全风险。作为 360 安全大脑 – 测绘云的核心系统，它将作为安全大脑的重要基础设施之一，成为连接现实世界与网络空间的桥梁。系统地址：quake.360.cn。

执行摘要

随着网络空间存储数据规模的急剧扩大，联网数据库发生数据泄露的条数和风险逐年增加。公开报道的在线数据泄露事件和数据库勒索事件屡创新高，在这背后的数据库安全问题频频被提及。那么，全球互联网上到底存在多少联网数据库？这些数据库类型和地域分布如何？又有多少比例的数据库存在大规模数据泄露风险？基于 360 自主研发的 Quake 网络空间测绘系统对全球 42 亿 IP 空间全面测绘的结果，我们对联网数据库进行了无害化探测发现和分析，时间从 2017 年至今，全面详细地展现了全球联网数据库分布特征和风险。研究报告的主要发现如下：

报告主要内容：

1、介绍了本报告所选取的数据库 MySQL、SqlServer、Oracle、PostgreSql、DB2、ElasticSearch、MongoDB、Memcache、Redis 和 CouchDB 十大数据库类型。我们根据数据库使用场景和分类分为了关系型数据库和非关系型数据库。探测发现，目前全球数据库联网数据库总量为 1500 万个，关系型数据库有 1400 万个，占总量的 93%，非关系型数据库有 109 万个，占总量的 7%。关系型数据库使用量要远大于非关系型数据的使用量。与 DB-Engines 的市场研究数据 Oracle 数据库排名第一不同的是，通过我们探测全网中 Oracle 仅为 157,157 个位于第六位。

2、对所选的数据库全球地理分布和利用希尔伯特曲线进行 IPV4 空间分布进行了分析。数据库全球地理分布主要集中在中美两国，在一些组织管理下的 IP 段，例如：RIPE NCC、ARIN 和 LACNIC 等数据库 IP 呈均匀分布，在另外一些，如 US-DOD 则不存在数据库 IP。在各大数据库中 MySQL 的使用量最多有 1 千多万。在各数

据库的全球分布中美国和中国都位于前两位，MySQL、PostgreSql、Redis、DB2 和 CouchDB 使用量最多的为美国，SqlServr、Oracle、MongoDB、ElasticSearch 和 Memcache 使用量最多的为中国。波兰在数据库总量位于第三，在 PostgreSQL 数据库使用量中位于第二位。同时把数据库探测得到的版本与各个版本发行日期和结束维护日期进行比较，发现全网仍有大量官方不在支持维护的数据库运行，如 Mysql 5.1 系列于 2013 年 12 月官方停止支持后，仍有 60 万个 5.1.73 和 16 万个 5.1.26 版本在使用。

3、对数据库存在泄露的情况进行分析，并提出数据库加固建议。通过分析发现，数据泄露仍是数据库安全的一大隐患，网中仍有超 8 万个数据库存在未授权访问漏洞。ElasticSearch 泄露数据量达到 3,402TB、MongoDB 泄露量为 611TB、Redis 泄露数据量为 10TB 和 Memcache 为 5.3TB。互联网约 30% 的 Memcache 数据库存未授权访问问题，ElasticSearch 存在未授权访问的数量占该数据库总量的 20% 左右。

4、将存在泄露的数据库 IP 进行归属地划分，我们发现在各个数据库泄露全球排名中，中国在数据库泄露数量排名全球第一具有近 4 万个数据库存在未授权访问漏洞，其中在 ElasticSearch、MongoDB 和 Redis 数据库类型中存在该漏洞的数量居全球第一，分别为 ElasticSearch 数据库 11,952 个、MongoDB 数据库 11,974 个和 Redis 7,127 个。南非在存在未授权访问的 Memcache 数据库排名第一，为 4890 个。

5、我们针对数据库勒索中出现的 BTC 地址做了统计，发现“1FYqD4YtPpcnHyyMiFFigG53s51dob6xx1”在勒索事件中出现次数最多高达 3,472 次，该数据反映出针对数据库的大规模、批量式勒索攻击依旧存在。

一、背景介绍

随着互联网的飞速发展，越来越多的设备接入互联网中形成网络空间。网络空间作为人类活动新的空间形态，是人和信息的共同载体[1]。数据库作为网络空间数据的承载基础设施，扮演了重要的角色，存储着人类在网络空间活动过程中产生的各级各类数据，有很多数据重要且敏感，涉及国家安全或个人隐私。这些数据一旦泄露将直接或间接造成重大经济损失。Quake 网络空间测绘系统是 360 网络安全响应中心（360-CERT）自主研发设计的全网空间测绘系统，能够对全球全量 IPv4、IPv6 地址进行持续性测绘工作，依托 360 全网海量大数据资源，具备全球网络空间测绘、监测能力。

2020 年 8 月 Ponemon Institute 和 IBM Security 联合发布的《2020 年数据泄露成本报告》[2]显示，2019 年 8 月至 2020 年 4 月间全球 17 个主要国家 / 地区 17 个行业共有 524 个组织报告发生了数据泄露事件，这些行业涵盖了医疗、金融、教育、能源、工业、通信等主要领域。数据泄露的平均总成本达到 386 万美元，该数字自 2014 年至今一直在 350 万到 400 万美元间波动，并没有显著改善。报告指出，当泄露数据条数在 100 万至 1000 万条时，平均总成本将达到 5000 万美元，而泄露数据条数超过 5000 万条时，平均总成本将飙升至 3.92 亿美元。数据泄露造成最严重的后果就是使组织的业务失效，由此遭受的损失占到平均总成本的 40%。下面 3 起 2020 年公开报道的事件显示了在线数据泄露的严峻形势。

背景事件一：美国超过 2 亿条人口信息泄露

2020 年 1 月 27 日 Comparitech 的安全研究员 Bob Diachenko 发现了一个未经任何权限验证的 ElasticSearch 数据库服务器暴露在网络上 [3]。该次数据泄露的样例数据如图 1-1 所示。

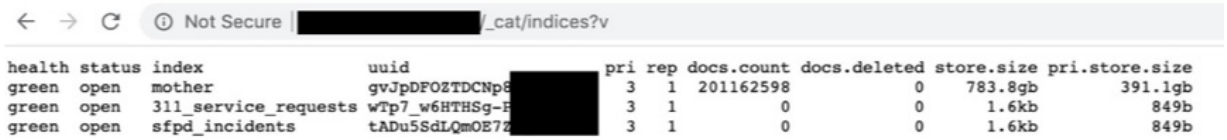


图 1-1 泄露数据样例

此次暴露的数据包括个人信息、人口统计信息和财产信息等共 201,162,598 条。这些数据在互联网中暴露时间长达一个多月，直到 2020 年 3 月 4 日该数据库服务器被关闭。在这一个多月时间里任何人都能够通过网络访问此数据库。由于泄露的数据包含详细的个人信息和财产信息，这些数据很可能被不法分子用来进行网络钓鱼等犯罪活动。

在本次事件中由于管理者未对数据库进行合理的权限验证，导致亿级美国个人信息数据在互联网公开暴露上长达一个多月时间。根据美国商务部下属的美国人口普查局数据显示 [4]，美国现阶段人口数量为 330,429,434 个，此次泄露的数据量超过美国人口总数的一半以上，且涉及到人口、财产等敏感信息。

背景事件二：英国印刷公司 Doxzoo 泄露 343GB 机密信息

2020 年 1 月 22 日 vpnMentor 网络安全团队发现了印刷公司 Doxzoo 的一个 Amazon S3 云存储服务器泄露了 343GB 大小的数据 [5] 其中包括个人护照（如图 1-2 左所示）信息、内部军事文件（如图 1-2 右所示）等敏感信息。根据 vpnMentor 描述，此次泄露有超过 10 万用户受

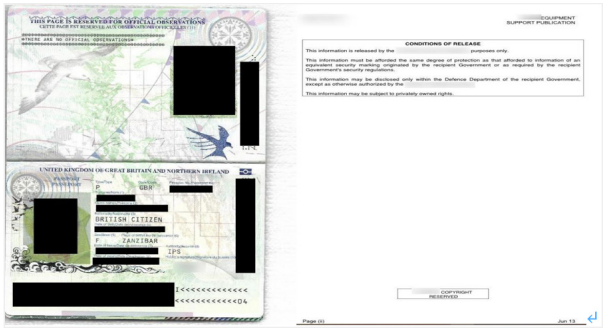


图 1-2 泄露的护照信息和泄露的军事材料

到影响，还涉及美国和英国的军事数据。该服务器于 2020 年 2 月 11 日关闭。由于未设置权限验证，在这之前任何人都可以通过互联网访问到该服务器。

在此事件中正是管理者未对亚马逊云存储服务进行正确的权限验证，导致大量敏感数据泄露，甚至包含了军事机密文件。本次事件警示，数据库安全可能威胁的不仅仅是个人，甚至可能涉及国家或者军事安全。

背景事件三：VPN 数据库泄露超过 2000 万用户日志

2020 年 7 月 5 日，vpnMentor 安全研究团队 Noam Rotem 发现了一台 ElasticSearch 服务器泄露了 1.207TB VPN 日志信息 [6]。数据泄露的数据摘要如表 2-1 所示。该服务器于 2020 年 7 月 15 日被关闭，其中泄露了包括明文密码、用户访问日志在内的 1.207TB 数据。一向以安全著称的 VPN 提供商，本次却由于低级的失误，造成了大量数据的泄露，此次泄露的数据也揭露了 No-log VPN“不记录用户访问日志”的谎言。

在此事件中正是管理者未对亚马逊云存储服务进行正确的权限验证，导致大量敏感数据泄露，甚至包含了军事机密文件。本次事件警示，数据库安全可能威胁的不仅仅是个人，甚至可能涉及国家或者军事安全。

表 2-1 VPN 日志泄露事件数据摘要

类别	描述
VPN 应用	UFO VPN, FAST VPN, Free VPN, Super VPN, Flash VPN, Secure VPN, Rabbit VPN
IP 所属	中国香港
行业	网络安全
数据量大小	1.207 TB
数量条数	1,083,997,361
影响人数	根据各个 VPN 应用声称的用户数量，超过 2000 万。
影响地理范围	全球
泄露的数据类型	活动日志、PII（姓名、邮件、家庭地址）、明文密码、比特币付款信息支持信息、个人设备信息、技术规格、账户信息、Paypal API 链接
数据库类型	ElasticSearch

二、方法

根据 DB-Engines 的市场研究数据 [7]，我们按照 2020 年数据库排名，选取了十个具有代表性的数据库进行了分析，如图 2-1 所示，反映了数据库的受欢迎程度。表 2-2 列出了我们本次分析的数据库类型，并按照数据库常见的使用场景和分类，划分为关系型和非关系型数据库。关系型数据库用来存储结构化数据，应用场景为数据关系性强，具有标准定义的项目。非关系型数据库适合存储非结构化数据，应用场景为一些数据量大，数据之间没有强关系的项目。

Rank			DBMS	Database Model	Score		
Oct 2020	Sep 2020	Oct 2019			Oct 2020	Sep 2020	Oct 2019
1.	1.	1.	Oracle	Relational, Multi-model	1368.77	-0.59	+12.89
2.	2.	2.	MySQL	Relational, Multi-model	1256.38	-7.87	-26.69
3.	3.	3.	Microsoft SQL Server	Relational, Multi-model	1043.12	-19.64	-51.60
4.	4.	4.	PostgreSQL	Relational, Multi-model	542.40	+0.12	+58.49
5.	5.	5.	MongoDB	Document, Multi-model	448.02	+1.54	+35.93
6.	6.	6.	IBM Db2	Relational, Multi-model	161.90	+0.66	-8.87
7.	↑ 8.	7.	Elasticsearch	Search engine, Multi-model	153.84	+3.35	+3.67
8.	↓ 7.	8.	Redis	Key-value, Multi-model	153.28	+1.43	+10.37
9.	9.	↑ 11.	SQLite	Relational	125.43	-1.25	+2.80
10.	10.	10.	Cassandra	Wide column	119.10	-0.08	-4.12

图 2-1 DB-Engines 数据库排名

本次探测和分析的数据全部取自 Quake 网络空间测绘系统无害扫描、去重得到，时间跨度从 2017 年 9 月到 2020 年 9 月。

表 2-2 本次分析的数据库

关系型数据库	非关系型数据库
MySQL	Redis
PostgreSql	ElasticSearch
SqlServer	MongoDB
DB2	Couchdb
Oracle	Memcache

三、数据分析

3、1 全球联网数据库地理分布概况

通过 Quake 平台，我们探测到全球数据库总量有 15,090,146 个，其中关系型数据库有 13,999,460 个，非关系型数据库有 1,090,686 个，如图 3-1 所示。

可以看到关系型数据库占总量的 92.7% ，非关系型数据库占总量的 7.3%，关系型数据库的数量远大于非关系型数据库。

图 3-1 数据库总量

将探测到数据库的服务 IP 进行归属地分类，可以得到数据库在全球国家的分布，如图 3-2 所示。

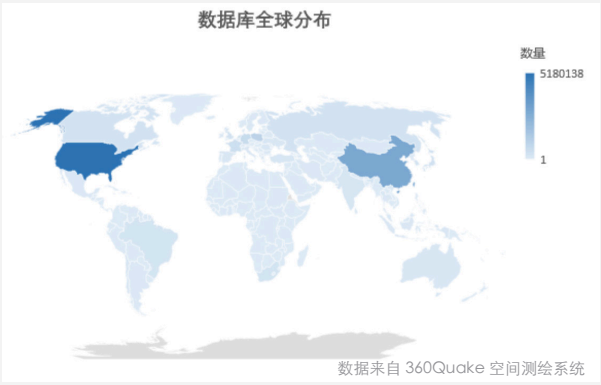


图 3-2 常见数据库全球分布

其中在各个国家中数量 TOP10 的如图 3-3 所示。结合全球分布来看数据库数量集中分布在北美洲、亚洲和欧洲。

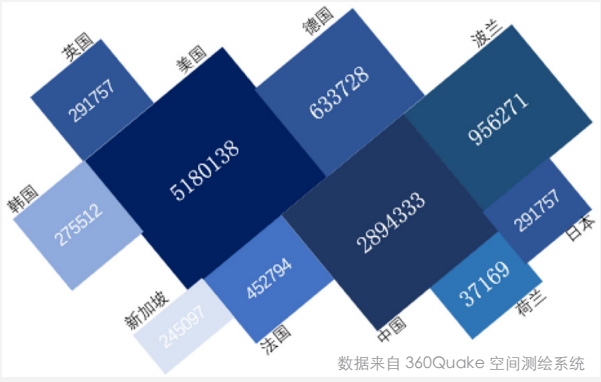


图 3-3 数据库资产量 TOP10

其中美国存在约 518 万个排名第一；中国 289 万个位于第二位；后续为波兰 956,271 个、德国 633,728 个、法国 452,794 个和荷兰 371,692，均为欧洲国家。在 TOP10 中令人意外的波兰排名第三，超过传统的德国和法国。

通过 Quake 平台，各类型数据库在公网暴露数量分布如图 3-4 所示。我们获得了 10,669,797 个运行 MySQL 服务的实例。PostgreSql 存在 1,515,560 个，SqlServer 有 1,282,744 个。探测数据和 DB-Engines 排名不同的是，MySQL 在互联网中暴露的数量远多于其它数据库，在 DB-Engines 排名第一的 Oracle 数据库在互联网上暴露数量为 157,157 个，位于第六位。

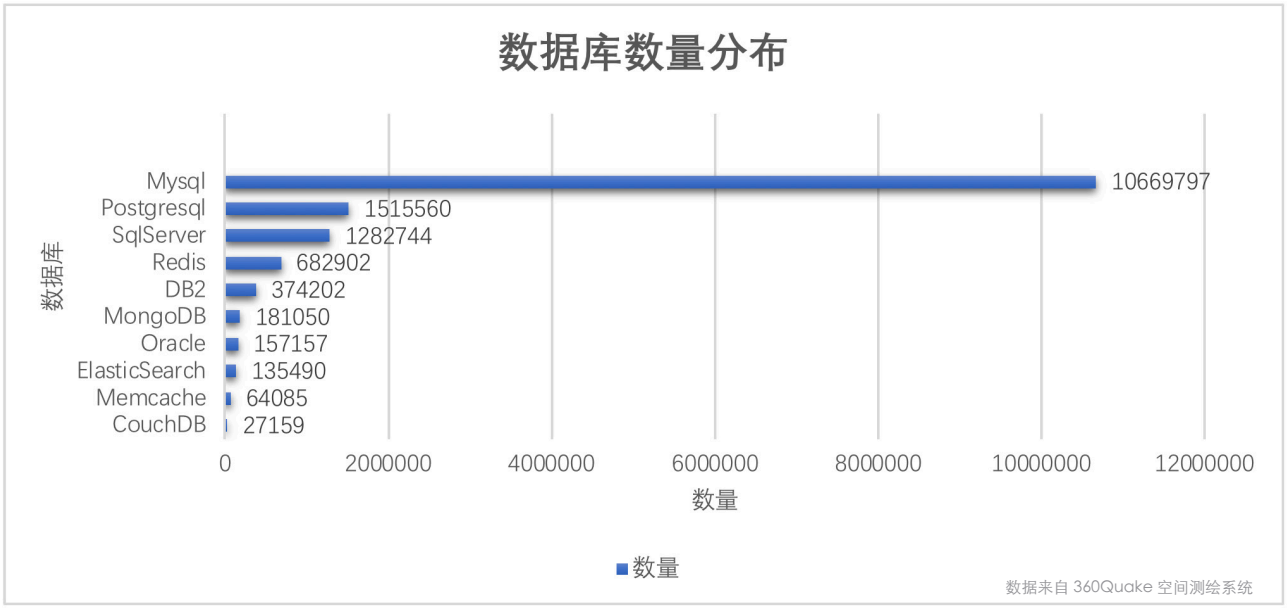


图 3-4 数据库数量分布

根据数据库在全球的分布，我们统计出了各个数据库使用量最多的国家（如表 3-1 所示）。在各个数据库使用量第一的国家中，MySQL、PostgreSql、Redis、DB2 和 CouchDB 使用量最多的为美国。SqlServr、Oracle、MongoDB、ElasticSearch 和 Memcache 使用量最多的为中国。

表 3-1 各个数据库使用量最多的国家

数据库	国家	数量
MySQL	美国	3,978,657
PostgreSql	美国	489,577
SqlServer	中国	304,049
Redis	美国	189,142
DB2	美国	103,303
Oracle	中国	57,500
MongoDB	中国	49,361
ElasticSearch	中国	43,705
Memcache	中国	19,001

为了对联网数据库在 IPV4 空间的分布一探究竟，我们使用希尔伯特曲线将一维的 IP 空间外推成二维空间，如图 3-5 所示。图中每个像素代表一个 C 段 IP 地址（/24），黑色像素代表该 C 段中数据库占比为 0%，蓝色到白色像素代表对应 C 段中数据库占比小于 50%，白色像素到红色像素之间代表占比大于 50%，红色像素代表占比 100%。由图 3-5 就可以看出，数据库在 IPV4 空间中总体上看并不是均匀分布的，而是星星点点成块分布。

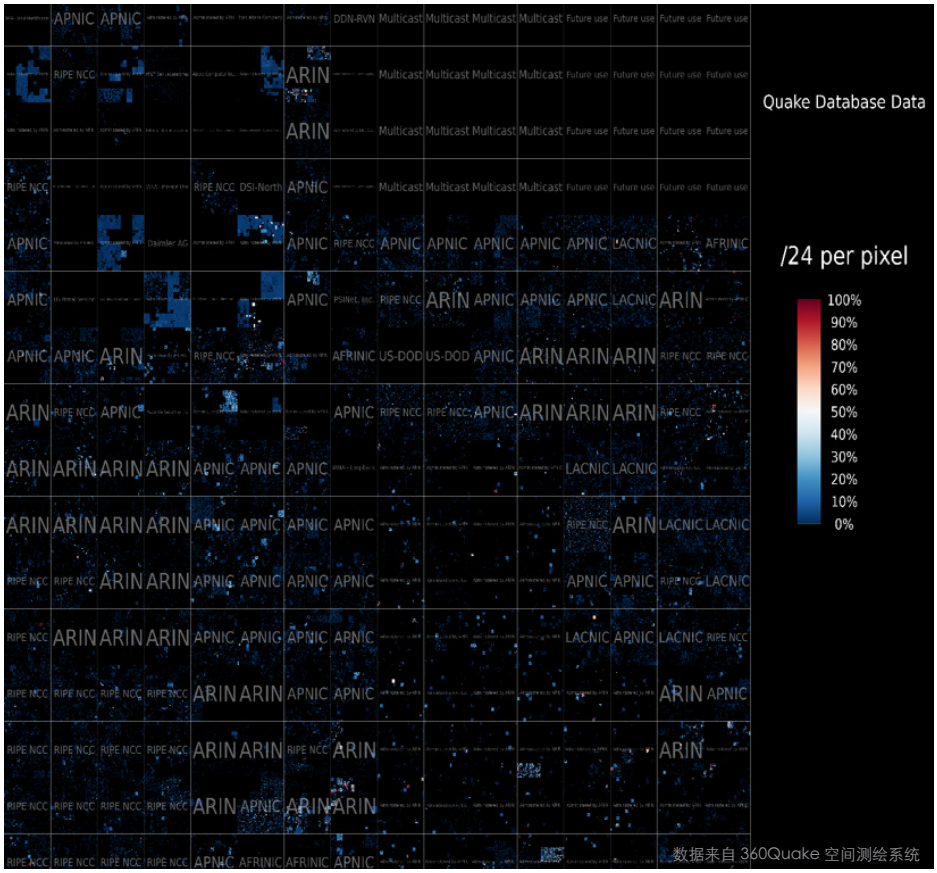


图 3-5 数据库资产在 IPV4 空间的分布

我们通过 IPV4 地址空间注册表 [8] 与我们探测得到的数据库 IP 进行匹配发现，在一些组织管理下的 IP 段，例如：RIPE NCC、ARIN 和 LACNIC 等数据库 IP 呈均匀分布，在另外一些，如 US-DOD 则不存在数据库 IP，如图 3-6 所示。

图 3-6 数据库资产在 IPV4 空间中部分组织的分布

3、2 主要数据库类型地理分布

3、2、1 MySQL

MySQL 是一个传统的关系型数据库，原由瑞典 MySQL AB 公司开发，2009 年被甲骨文公司(Oracle) 收购，成为 Oracle 旗下产品。MySQL 由于性能高、成本低和高可靠性被广范应用。在被甲骨文收购后，MySQL 创始人麦克尔 · 维德纽斯以 MySQL 为基础，成立分支计划 MariaDB。本次探测 MySQL 的数据量为 MySQL 和 MariaDB 数据总和。

我们探测得到的 MySQL 总量为 10,669,797 个 ,MySQL 全球分布如图 3-7 所示。在 MySQL 全球分布的国家中，美国使用量最多，有 3,978,657 个，中国使用量 2,203,568 个位于第二位，德国拥有 477,824 个位于第三。第四位到第十位的分别是波兰 405,029 个、法国 380,792 个、荷兰 254,156 个、英国 214,878 个、南非 211,435 个、日本 671976 个和俄罗斯联邦 192,674 个，如表 3-2 所示。

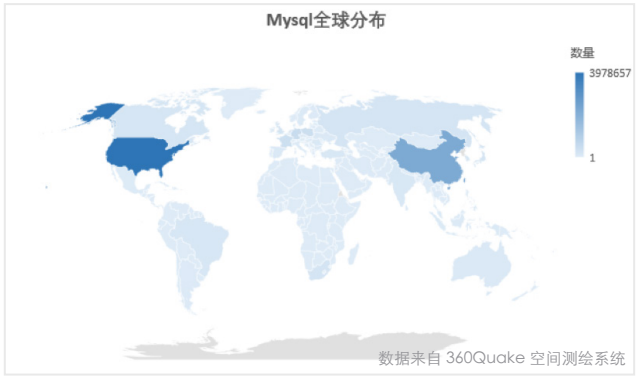


图 3-7 MySQL 全球分布

表 3-2 MySQL 分布国家 TOP 10

国家	数量
美国	3,978,657
中国	2,203,568
德国	477,824
波兰	405,029
法国	380,792
荷兰	254,156
英国	214,878
南非	211,435
日本	200,453
俄罗斯联邦	192,674

在 MySQL 的各个版本中，版本 5.6.41 使用量最大有 883,310 个，版本 5.1.73 使用量为 60,2291 位于第二位，如图 3-8 所示。

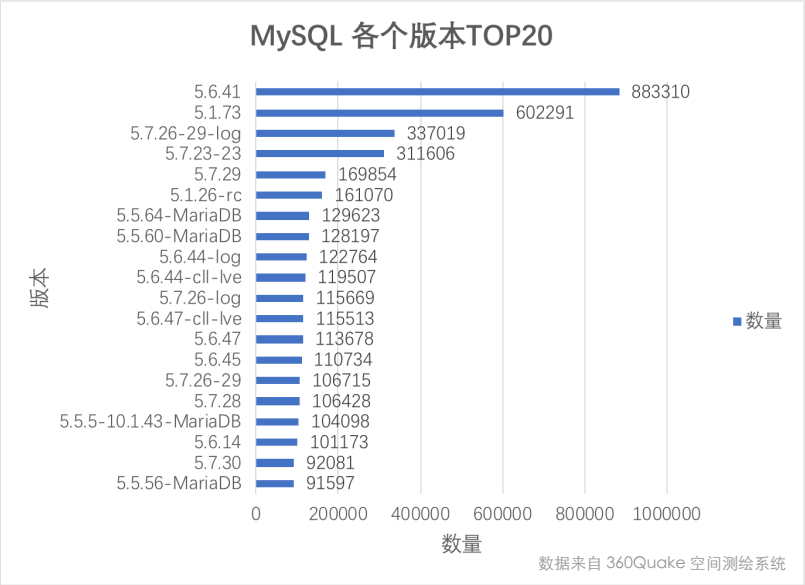


图 3-8 MySQL 各个版本 TOP20

表 3-3 罗列了 MySQL 各个版本的发布日期和终止支持的日期。在 MySQL 版本 TOP20 中发现，仍有大量官方不再维护的老版本的 MySQL 在互联网中存在，例如 MySQL 5.1.73 数量为 602,291 个，位于 MySQL 各个版本的第二位，距离官方终止支持的日期过去了近 7 年。

表 3-3 MySQL 各个版本发布时间

版本	发布日期	终止支持的日期
MySQL 5.1	2008 年 11 月 14 日	2013 年 12 月
MySQL 5.5	2010 年 3 月 12 日	2018 年 12 月
MySQL 5.6	2013 年 5 月 21 日	2021 年 2 月
MySQL 5.7	2015 年 10 月 21 日	2023 年 10 月
MySQL 8.0	2018 年 4 月 19 日	2026 年 4 月

3、2、2 ElasticSearch

ElasticSearch 是一个基于 Luncene 库的搜索引擎，基于 Java 语言开发的一个流行的企业级搜索引擎。2010 年 Shay Banon 发布了 ElasticSearch 的第一个版本。ElasticSearch 常用于云计算中，能够达到实时搜索、稳定、可靠、快速，安装使用方便。ElasticSearch 也由于未授权访问漏洞，常常被爆出数据泄露事件。因此，对 ElasticSearch 进行测绘非常有意义。

我们在互联网中共发现了 135,490 个开启 ElasticSearch 服务的实例，全球分布如图 3-9 所示。

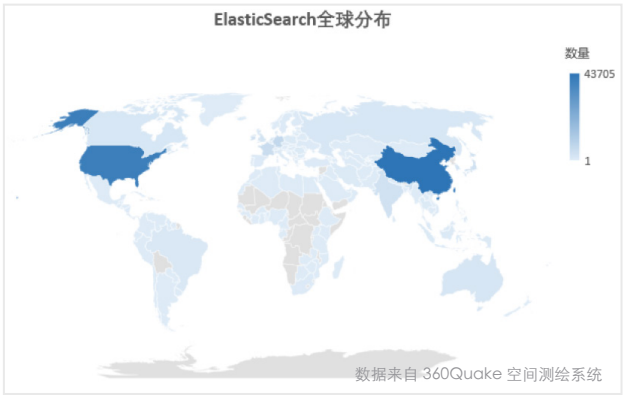


图 3-9 ElasticSearch 全球分布

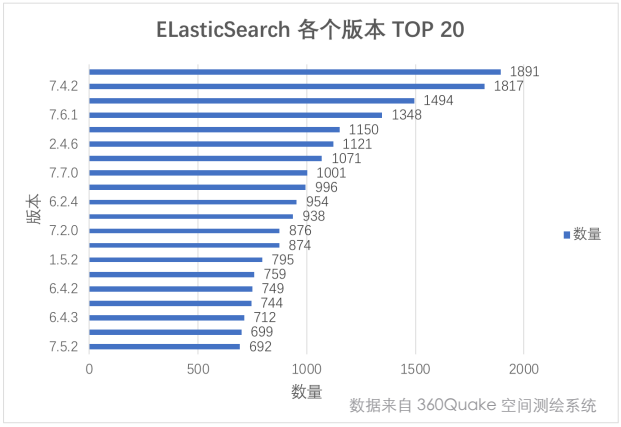


图 3-10 ElasticSearch 使用版本 TOP20

表 3-4 ElasticSearch 分布国家 TOP 10

国家	数量
中国	43,705
美国	40,163
德国	6,605
法国	5,074
新加坡	5,060
爱尔兰	4,094
日本	3,445
印度	3,328
荷兰	3,190
韩国	2,629

ElasticSearch 全球分布 TOP10 如表 3-4 所示，中国 ElasticSearch 使用量最多有 43,705 个；其次为美国 40,163 个，和中国数量较为接近；德国 6,605 个排名第三。我们还统计了 ElasticSearch 各个版本的数量如图 3-10 所示。ElasticSearch 目前使用的最多的版本是 7.6.2，有 1,891 个，并且通过图 3-9 还能发现，目前仍有不少老版本的 ElasticSearch 在使用，如 2.4.6 版本有 1,121 个，1.5.2 版本有 759 个。

3、2、3 Redis

Redis（Remote Dictionary Server），即远程字典服务，是一个开源、支持网络、日志型的 Key-Value 数据库，并提供多种语言的 API。

我们探测得到的 Redis 有 682,902 个，全球分布如图 3-11 所示。Redis 的全球分布和 ElasticSearch 的很类似，主要集中在欧美和中国等地，其中分布 TOP10 如表 3-5 所示。

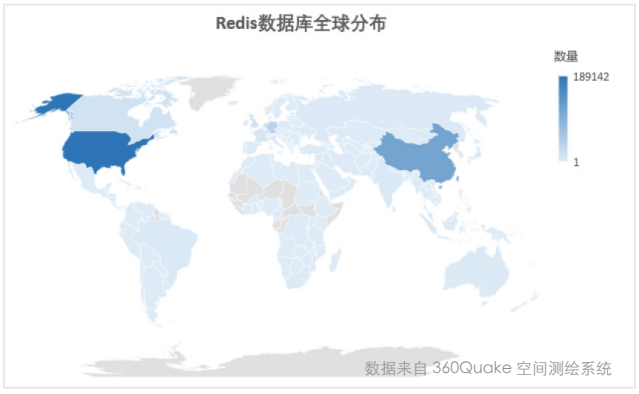


图 3-11 Redis 全球分布

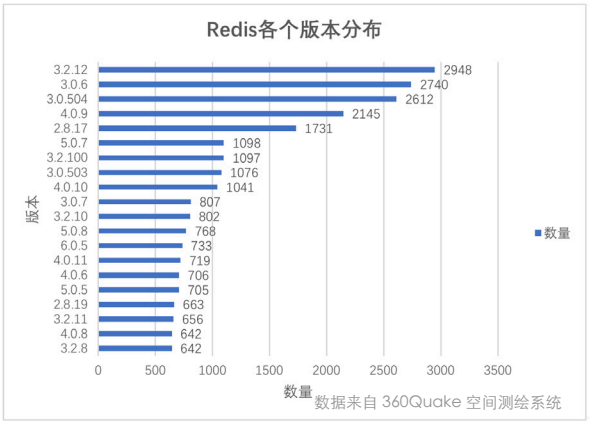


图 3-12 Redis 使用版本 TOP20

表 3-5 Redis 分布国家 / 地区 TOP 10

国家	数量
美国	189,142
中国	114,239
荷兰	37,235
德国	37,001
INCAPSULA.COM	32,401
Europe Regions	25,995
英国	17,337
加拿大	13,604
亚太	10,759
法国	8,163

美国的 Redis 使用量最多有 189,142 个，中国有 114,239 个位列第二。同时我们也列出了 Redis 各个版本的统计，Redis 版本 TOP20 如图 3-12 所示，从图中可以看到 3.2.12 版本使用量最多，有 2948 个。

3、2、4 MongoDB

MongoDB 是一种面向文档的数据库，由 C++ 语言编写而成，初始版本于 2009 年 2 月发布。MongoDB 具有强大的查询语言，可以实现与关系型数据库表单查询类似的大部分功能，支持对数据建立索引。通过探测，我们发现 MongoDB 全网共有 181,050 个，全球分布如图 3-13 所示。

表 3-6 罗列了 MongoDB 使用数量 TOP10 的国家。美国和中国的使用量相接近，分别为 49,361 个和 48,372 个。其它国家和前两名差距较大，日本仅有 21,068 个位于第三位。其后，德国、新加坡和法国的使用量较为接近分别为 9,699 个、7,574 个和 6,065 个。印度和荷兰的使用量较为接近分别为 4,722 个和 4,360 个。爱尔兰和英国 MongoDB 使用量较为接近分别为 3,998 个和 3,323 个。

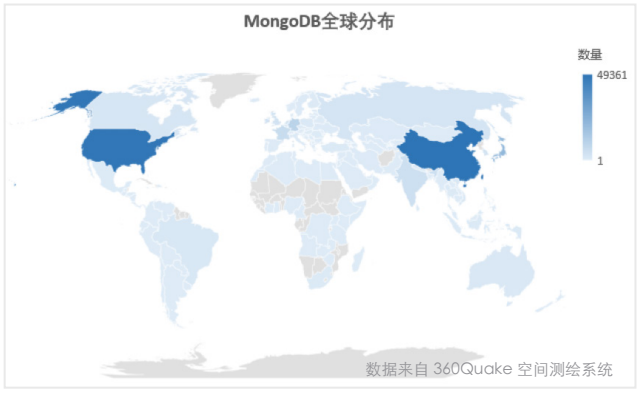


图 3-13 MongoDB 全球分布

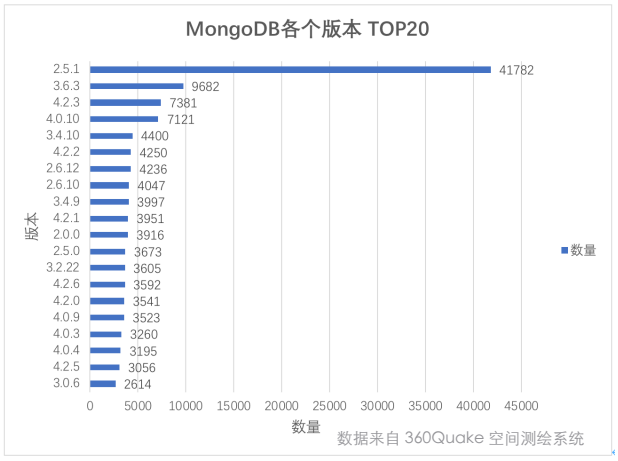


图 3-14 MongoDB 各个版本 TOP20

表 3-6 MongoDB 分布国家 TOP 10

国家	数量
中国	49,361
美国	48,372
日本	21,068
德国	9,699
新加坡	7,574
法国	6,065
印度	4,722
荷兰	4,360
爱尔兰	3,998
英国	3,323

同样的，我们对 MongoDB 的版本分布做了统计，如图 3-14 所示。其中，2.5.1 版本远多于其它版本，有 41,782 个。根据 MongoDB 的各个版本的发行日期和终止支持日期 [9]，可以发现仍有大量的老版本 MongoDB 在互联网上使用。

3、2、5 Memcache

Memcache 是一套分布式的高速缓存系统，由 LiveJournal 的 Brad Fitzpatrick 开发。我们探测得到的 Memcache 有 64，085 个，Memcache 的全球分布，如图 3-15 所示。

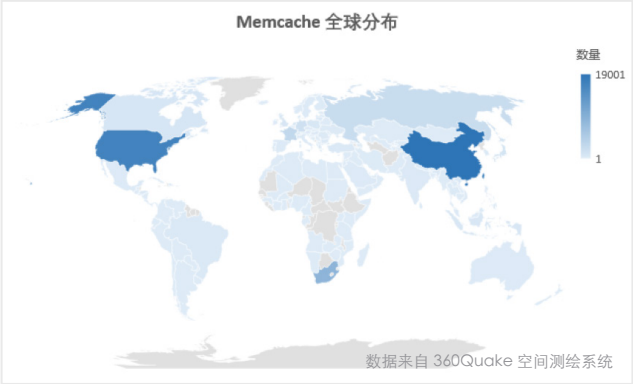


图 3-15 Memcache 全球分布

根据表 3-7 所示，在 Memcache 全球分布的国家中，中国使用量最多，有 19,001 个；美国使用量 16,693 个位于第二位；南非拥有 8,686 个位于第三。之后数量急剧减少，从法国 2,973 个开始依次为俄罗斯联邦 2,274 个、德国 1,604 个、日本 1,037 个、荷兰 863 个、英国 748 个和加拿大 727 个。

表 3-7 Memcache 分布国家 TOP 10

国家	数量
中国	19,001
美国	16,693
南非	8,686
法国	2,973
俄罗斯联邦	2,274
德国	1,604
日本	1,037
荷兰	863
英国	748
加拿大	727

3、2、6 PostgreSql

PostgreSql 是一种特性非常齐全的关系型数据库，是以加州大学计算机系开发的 POSTGRES 4.2 版本为基础的对象关系型数据库。

PostgreSql 在互联网中共有 1,515,560 个，PostgreSql 的全球分布如图 3-16 所示。如表 3-8 所示，在 PostgreSql 全球分布的国家中，美国使用量最多，有 489,577 个；波兰使用量 380,794 个位于第二位；中国拥有 77,241 个位于第三。之后从德国 69,814 个开始依次为智利 59,614 个、巴西 43,422 个、韩国 36,647 个、日本 34,537 个、法国 27,541 个和爱尔兰 25,385 个。

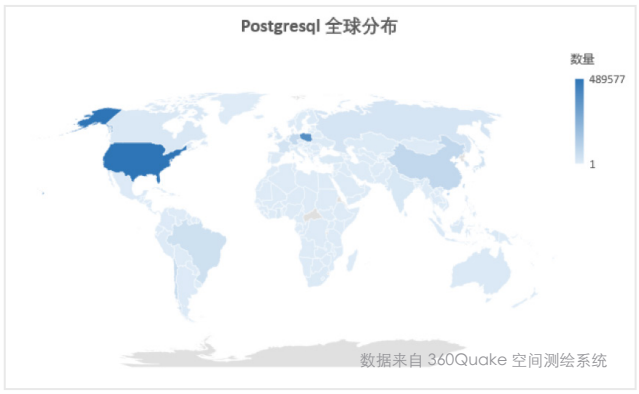


图 3-16 PostgreSql 全球分布

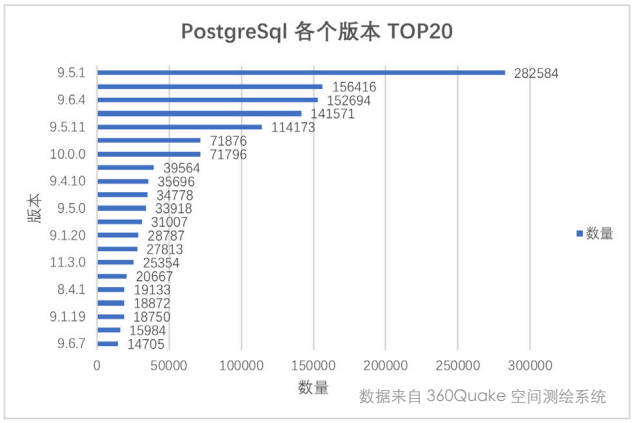


图 3-17 PostgreSql 各个版本 TOP20

表 3-8 PostgreSql 分布国家 TOP 10

国家	数量
美国	489,577
波兰	380,794
中国	77,241
德国	69,814
智利	59,614
巴西	43,422
韩国	36,647
日本	34,537
法国	27,541
爱尔兰	25,385

如图 3-17 所示，在 PostgreSql 的各个版本中，版本 9.5.1 使用量最多有 282,584 个。

3、2、7 SqlServer

SqlServe 是由 Microsoft 开发和推广的，在全网中共有 1,282,744 个，SqlServer 的全球分布如图 3-18 所示。表 3-9 罗列了 SqlServer 在全球主要国家的分布。中国使用量最多，有 304,049 个；美国使用量 282,948 个位于第二位；波兰拥有 166,253 个位于第三。之后从韩国 41,672 个开始依次为印度 36,565 个、土耳其 30,846 个、巴西 29,946 个、德国 20,727 个、越南 19,069 个和智利 18,018 个。

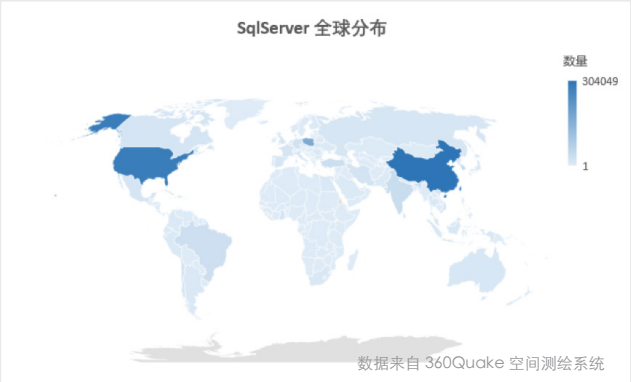


图 3-18 SqlServer 全球分布

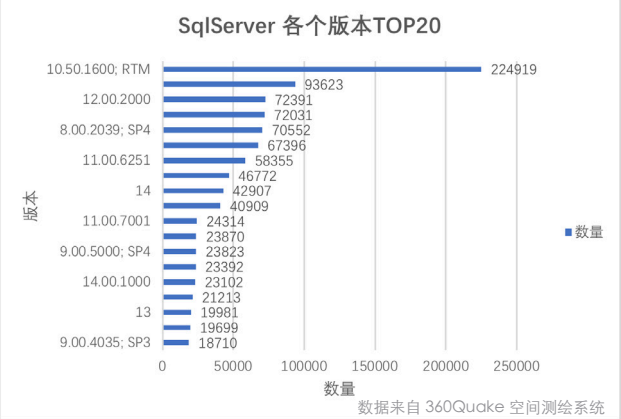


表 3-9 SqlServer 分布国家 TOP 10

表 3-9 SqlServer 分布国家 TOP 10

国家	数量
中国	304,049
美国	282,948
波兰	166,253
韩国	41,672
印度	36,565
土耳其	30,846
巴西	29,946
德国	20,727
越南	19,069
智利	18,018

图 3-19 显示在 SqlServer 的各个版本中，版本 10.50 使用量最多有 224,919 个，远远多于其它版本。

3、3 主要数据库风险分析

在 3.2 节中提到，通过 Quake 平台发现 Redis 总共有 682,902 个开放在公网的实例，ElasticSearch 和 MongoDB 分别有 135,490 和 181,050 个，而 Memcache 实例有 64,085 个。过往的一些严重的数据泄露事件常常与 ElasticSearch 和 MongoDB 数据库相关，这些数据库通常存储了大量数据，但管理者往往会忽视对这些数据库进行权限验证（未授权访问漏洞），导致数据泄露。在得知这些数据库在互联网上暴露的数量后，我们还想知道，到底还有多少数据库存在未授权访问漏洞，还有多大的数据暴露在互联网上。

我们把测绘数据中存在未授权访问漏洞的数据库与数据库总量做对比，如图 3-20 和表 3-10 所示，Redis 中存在问题的实例占总数的比例为 2.15%，有 14,704 个。ElasticSearch、MongoDB 和 Memcache 中存在该问题的实例占总量比例分别为 20.26%、12.28% 和 31.80%。

表 3-10 数据库存在未授权访问漏洞数量占比

数据库类型	具有未授权访问漏洞占比
Memcache	31.80%
ElasticSearch	20.26%
MongoDB	12.28%
Redis	2.15%

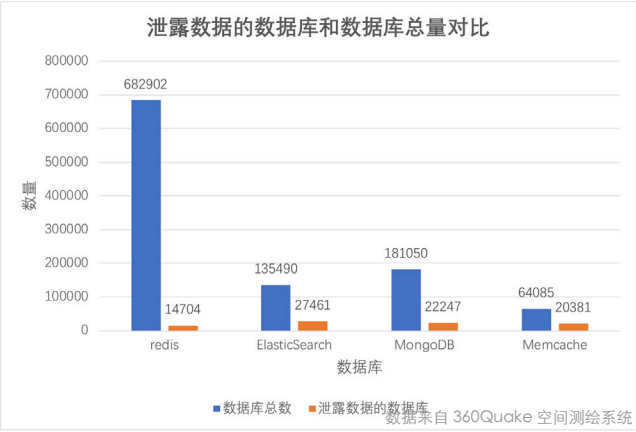


图 3-20 未授权漏洞的数据库与数据库总量对比

需要引起重视的是，仍然有 822 个独立 IP 在 UDP11211 端口上运行着 Memcache。而早在 2017 年 360 信息安全部 0kee Team 就发现了利用此配置进行的 Memcache DRDoS 攻击，并在 PoC 2017 会议上做了公开报告进行预警 [10]。

在多次此类攻击事件发生后，2018 年 360 信息安全部 0kee Team、360 网络安全研究院、360-CERT 还共同发布了技术博客深入分析了该漏洞 [11]。

在明确泄露实体数量后，我们根据如下字段进行分析，分别为：Redis 中的 used_memory、Memcache 中的 bytes、MongoDB 中的 datasize 和 ElasticSearch 的 size 字段。我们利用这些字段数据，做了进一步的探测，将各个实体泄露数量相加得出各个数据库可能泄露的总量，如图 3-21 所示。ElasticSearch 泄露数据量达到 3,402TB、MongoDB 泄露量为 611TB、Redis 泄露数据量为 10TB 和 Memcache 为 5.3TB。可以看到 ElasticSearch 和 MongoDB 泄露的数据量远大于 Redis 和 Memcache，这和 ElasticSearch 常用于大数据搜索有关。

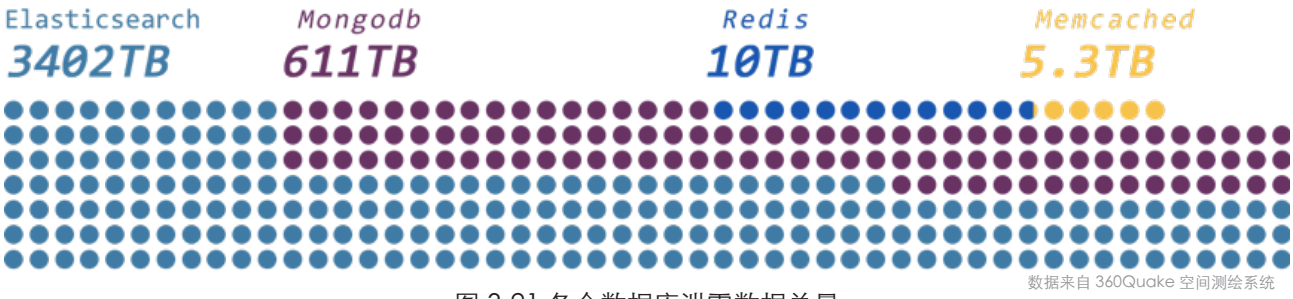


图 3-21 各个数据库泄露数据总量

将泄露数据的数据库 IP 进行归属统计，可以得到数据库泄露的全球分布，如图 3-22 所示。根据表 3-11 数据库泄露国家 TOP10 所示，中国有 35,095 个 IP 可能存在数据泄露，位于第一，美国有 17,830 个位于第二。ElasticSearch、Memcache、MongoDB 和 Redis 泄露数量最多的国家如表 3-12 所示。可以看出除 Memcache 外，中国在其它三个数据库中可能泄露数量是最多的。

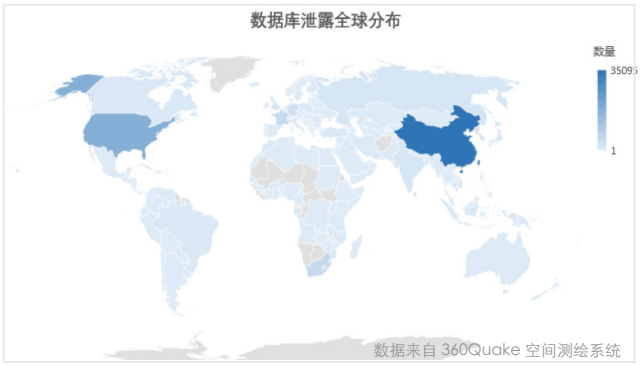


图 3-22 数据库泄露全球分布

表 3-12 各个数据库泄露数量最多的国家

数据库类型	国家	具有泄露风险的数据库数量
ELasticSearch	中国	11,952
MongoDB	中国	11,974
Redis	中国	7,127
Memcache	南非	4,890

表 3-11 数据库泄露国家 TOP10

国家	数量
中国	35,095
美国	17,830
南非	5,030
法国	4,687
德国	3,000
新加坡	1,830
俄罗斯联邦	1,574
印度	1,526
韩国	1,492
日本	1,420

在我们统计了存在数据泄露的数据库后，又统计了有关数据库被勒索的情况。在我们的探测中发现，存在数据泄露（未授权访问漏洞）的数据库中，会存在一些如：“READ_ME_TO_RECOVER_YOUR_DATA”、“have_7days_to_contact_us”等勒索字样，一个勒索实例如图 3-23 所示。一些攻击者会利用未授权访问漏洞，窃取数据库数据并留下相关比特币地址进行勒索。我们统计了出现次数最多的数据库勒索比特币地址，如表 3-13 所示。

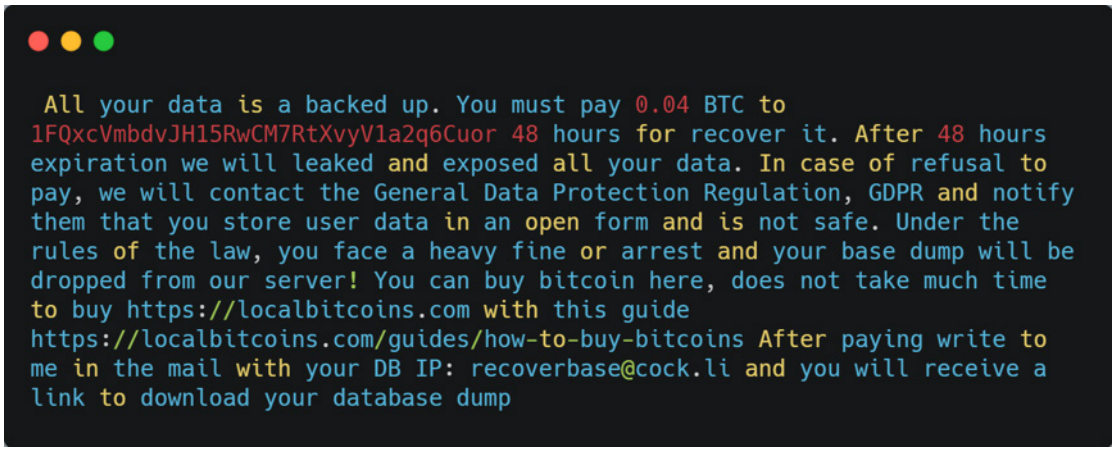


图 3-23 一个数据库被勒索实例

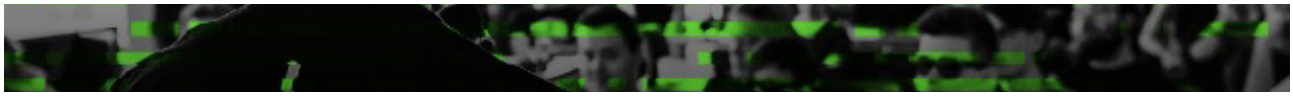
表 3-13 数据库勒索中 BTC 地址 TOP10

地址	数量
1FYqD4YtPpcnHyyMiFFigG53s51dob6xx1	3,472
1FQxcVmbdvJH15RwCM7RtXvyV1a2q6Cuor	3,035
1FQxcVmbdvJH15RwCM7RtXvyV1a2q6Cuor	1,966
1FLjhopErMQrNMSgDQQe7kcZUYcyGcxX8g	1,337
1JL7nX8AFhF5dCpqVNfxCCtZc5vVJwJTPN	1,288
1FbZiNH786rxDFtCds9vtLinHhuySV9Kmi	55
1FeNQxEJx2tz4Wv3njJmBT7sCHuDhDhcAi	16
1LSMdtnKDjbWrtQJmEtXPJhha9KnRXGYhW	16
1MPeDWguA9uqgvv4SKmcWQ47rcxGB6EkiT	14
17JbK3n4Emn3w7Q6tMA9bPcsYqNELv3j1b	12

3、4 主要数据库风险分析

通过上述分析结果，我们发现仍然有大量的联网数据库存在基础安全配置缺失的问题。因此，我们汇总了数据库通用加固建议，方便读者进行数据库安全的基本检查：

- 数据库应配置为始终需要进行身份验证。
- 分配数据库用户权限应遵守最小权限原则 [12]，既仅具有应用程序所需的最小权限。
- 不使用 root、sa 或 sys 等内置的或默认的账号。
- 不对数据库实例添加账户管理权限。
- 设置白名单，仅白名单内的主机可具有对数据库访问权限。
- 仅授予用户对所需特定数据库的访问权限。
- 开发环境、测试环境和生产环境应该使用单独的数据库和账户。



我们还列举了各个数据库官方安全建议，方便查阅：

1. MySQL 安全指南：https://dev.MySQL.com/doc/refman/8.0/en/security-guidelines.html
2. ostgreSql 安全指南：https://www.PostgreSql.org/docs/12/runtime.html
3. SqlServer 安全指南：https://docs.microsoft.com/en-us/sql/relational-databases/security/securing-sql-server
4. Oracle 安全指南：https://docs.oracle.com/en/database/oracle/oracle-database/12.2/dbseg/index.html
5. DB2 安全指南：https://www.ibm.com/support/knowledgecenter/SSEPEK_11.0.0/pdf/db2z_11_secabook.pdf
6. ElasticSearch 安全指南：https://www.elastic.co/guide/en/elasticsearch/reference/current/configuring-security.html
7. MongoDB 安全指南：https://docs.mongodb.com/manual/administration/security-checklist/
8. Redis 安全指南：https://redis.io/topics/security
9. Couchdb 安全指南：https://guide.couchdb.org/draft/security.html
10. Memcache 安全指南：https://blog.couchbase.com/memcached-security-best-practices

四、结论

数据库是互联网服务组成的重要应用。也是 360 Quake 网络空间测绘系统重点关注的专题方向。通过 Quake 平台，我们探测到全球数据库总量有 15,090,146 个，其中关系型数据库有 13,999,460 个，非关系型数据库有 1,090,686 个，本研究报告基于 Quake 对全网空间扫描的结果，通过对全球互联网最受欢迎的前几种数据库进行深入探测和数据分析，描绘了其全球国家分布状况。同时我们对数据库数据泄露风险也进行了评估。这是第一个对全网数据库详尽研究的分析报告，极具价值。我们的分析显示，当前全球主要国家数据库泄露风险不容乐观，数据库防护亟待加强。主要结论如下：

1. 探测选取了全球使用广泛的 MySQL、SqlServer、Oracle、PostgreSql、DB2、ElasticSearch、MongoDB、Memcache、Redis 和 CouchDB 十大数据库类型，前 5 个属于关系型数据库，后 5 个属于非关系型数据。探测发现关系型数据库的使用量远比非关系型数据库的使用量多，全球联网数据库超 93% 都是关系型数据库。
2. 各数据库类型在全球的国家分布中，美国和中国都位于前两位。值得注意的是位于波兰的联网数据库总量位于全球第三，而在 PostgreSql 数据库中排名第二。
3. 探测显示在互联网中仍存在大量的未授权访问的数据库，这些数据库包含了大量的数据。较为严重的数据库类型是 Memcache，超过 30% 的实例都存在未授权访问漏洞；而 ElasticSearch 类型超过 20% 存在该漏洞。
4. 探测显示我国还有大量数据库存在数据泄露的风险，通过对 ElasticSearch、MongoDB、Redis 和 Memcache 探测发现仍有近 4 万个数据库存在未授权访问漏洞。其中 ElasticSearch、MongoDB 和 Redis 数据库类型中存在该漏洞的数量居全球第一。

五、术语表

[1] 数据库：存储数据的仓库，是长期存放在计算机内、有组织、可共享的大量数据的集合。数据库中的数据按照一定数据模型组织、描述和存储，具有较小的冗余度，较高的独立性和易扩展性，并为各种用户共享，即数据库有永久存储、有知识和可共享三个基本特点。

[2] 关系型数据库：指采用了关系模型来组织数据的数据库，其以行和列的形式存储数据，以便于用户理解，关系型数据库这一系列的行和列被称为表，一组表组成了数据库。

[3] 非关系型数据库：区别于关系数据库，它们不保证关系数据的 ACID 特性。非关系型数据库都具有非常高的读写性能，无须事先为要存储的数据建立字段，随时可以存储自定义的数据格式。

[4] 空间测绘：网络空间资源测绘是对各类网络空间资源及其属性进行探测、融合分析和绘制。

[5] 希尔伯特曲线：是由大卫·希尔伯特在 1891 年提出的一条能充满整个平面正方形的曲线。本文中将开发数据库服务的 IPV4 地址，通过该方式进行展现。

六、参考

[1] 网络空间地理学的理论基础与技术路径 [J]. 高春东 , 郭启全 , 江东 , 王振波 , 方创琳 , 郝蒙蒙 . 地理学报 . 2019(09)

[2] https://www.ibm.com/downloads/cas/BK0BB0V1

[3] https://www.comparitech.com/blog/vpn-privacy/200-million-us-database-leaked/

[4] https://www.census.gov/popclock/

[5] https://www.vpnmentor.com/blog/report-doxzoo-leak/

[6] https://www.vpnmentor.com/blog/report-free-vpns-leak/

[7] https://db-engines.com/en/ranking

[8] https://www.iana.org/assignments/ipv4-address-space/ipv4-address-space.xhtml

[9] https://docs.mongodb.com/manual/release-notes/

[10] http://powerofcommunity.net/poc2017/shengbao.pdf

[11] https://blog.netlab.360.com/memcache-ddos-a-little-bit-more-en/

[12] https://zh.wikipedia.org/wiki/%E6%9C%80%E5%B0%8F%E6%9D%83%E9%99%90%E5%8E%9F%E5%88%99